# НАМАЛЯВАНЕ НА РЕКЛАМНИТЕ ИЗМАМИ В ДИГИТАЛНИЯ МАРКЕТИНГ С ПОМОЩТА НА ИЗКУСТВЕН ИНТЕЛЕКТ

**Ивона Велкова**

Асистент, доктор, катедра „Информационни технологии и комуникации“, УНСС
*e-mail:* ivonavelkova@unwe.bg

**Резюме**

*Непрекъснатата дигитализация на глобалната икономика поставя началото на свързаност и възможности за бизнеса да се ангажира с потребителите. Въпреки това, тя също така дове жда и до различни заплахи в онлайн пространството. Една от тези заплахи е рекламната изм ама, която застрашава целостта и ефективността на онлайн рекламните кампании в все по-дигитализираната икономика. Изкуственият интелект (ИИ), оборудван с възможности за ан ализ на данни в реално време и усъвършенствани алгоритми за идентифициране на модели, пре длага възможности за ефективно реагиране срещу този вид атаки. Чрез анализа на множест ва от данни, ИИ може да разграничи истинските потребителски взаимодействия в дигитален ото пространство от измамните, да маркира аномалии и да се адаптира бързо към развиващ ите се измамни тактики. Този документ се представя подход за решение за смекчаване на рек ламните измами, със специфичен акцент върху тяхното въздействие и ролята на ИИ в справя нето с това предизвикателство и насърчава създаването на по-сигурна и устойчива дигитален а икономика.*

**Ключови думи:** изкуствен интелект, заплахи, дигитален маркетинг

# REDUCING ADVERTISING FRAUD IN DIGITAL MARKETING WITH ARTIFICIAL INTELLIGENCE

**Ivona Velkova**

**Abstract**

*The continuous digitization of the global economy brings about connectivity and opportunities for bus inesses to engage with consumers. However, it also introduces various threats in the online space. On e of these threats is ad fraud, which jeopardizes the integrity and effectiveness of online advertising ca mpaigns in an increasingly digitized economy. Artificial intelligence (AI), equipped with real-time dat a analysis capabilities and advanced algorithms for pattern recognition, offers effective ways to count eract such attacks. Through the analysis of extensive datasets, AI can distinguish genuine user interac tions in the digital realm from fraudulent ones, identify anomalies, and adapt rapidly to evolving fraud ulent tactics. This document presents an approach to mitigating ad fraud, with a specific focus on its i mpact and the role of AI in addressing this challenge, promoting the creation of a more secure and res ilient digital economy.*

**Keywords:** artificial intelligence, threats, digital marketing
**JEL:** M31, O33

## Въведение

Дигитализацията на глобалната икономика представлява промяна в начина, по който се провеждат икономическите дейности, обменът на данни и достъпът до пазарите. През последните десетилетия, бизнес средата свидетелства за изключителна трансформация в начина, по който компаниите рекламират своите продукти и услуги. Тази промяна се движи от

напредъка в технологиите като Интернет, изкуствен интелект (ИИ), разпространението на мобилни устройства и растежа на платформите за електронна търговия. Процесът на дигитализация довежда до повишена свързаност и взаимозависимост между различните индустрии и потребителите в глобален мащаб [1], [2].

Интегрирането на технологиите не само фундаментално трансформира сектори като икономиката, здравеопазването и образованието, но също така очертава ключови роли за редица направления, включително и дигиталния маркетинг. Последният се използва в икономиката като незаменим компонент на съвременните маркетингови стратегии, използвани от компаниите по целия свят. За да може бизнесът да достигне до клиентите той използва дигиталната реклама, която е ключов компонент на дигиталния маркетинг. Тя предлага на бизнеса възможността да се свързва с целевата си аудитория в реално време, предоставяйки персонализирано съдържание и решения. Дигиталният маркетинг включва широк спектър от онлайн дейности, насочени към достигане и ангажиране с целеви аудитории чрез дигитални канали като уебсайтове, социални мрежи, електронна поща и различни форми на онлайн съдържание. В резултат на това онлайн рекламата се превръща в неразделна част от маркетинговите кампании и ключов двигател на приходите в дигиталната икономика [3], [4].

Въпреки че използването на дигиталната реклама донася множество предимства за бизнеса, то внася и нови предизвикателства като част от тях са трафик от ботове (интернет-трафик, който е генериран от автоматизирани програми), измами с данни, рекламни измами и др. Последните обхващат различни практики, целящи манипулиране на онлайн рекламни кампании, като например фалшиви кликове, измами при показвания и стекинг на реклами [5]. Рекламните измами имат вредни ефекти както върху бизнесите, така и върху потребителите. За бизнесите, те довеждат до финансови загуби, тъй като плащат за измамни взаимодействия, които не предоставят истинска стойност. Освен това, този вид измами развалят репутацията на марката, тъй като потребителите могат да я асоциират с измамни дейности. Проучване на Juniper на установява, че 22% представляват значителна част от общите разходи за онлайн реклама на компаниите, които са отрицателно повлияни от рекламни измами, което допълнително подчертава значителната финансова тежест, която налага на бизнеса [6]. За потребителите, рекламните измами могат да доведат до лошо онлайн преживяване, излагане на потенциално вредно съдържание и намаляване на доверието в дигиталната реклама.

За да се предотвратят тези предизвикателства, използването на напреднали технологии като изкуствен интелект (ИИ) се явява ключова стратегия. ИИ-базираните решения предоставят възможности за по-ефективно и активно откриване и предотвратяване на измамните дейности, тъй като имат способността да анализират огромни количества данни в реално време. Това означава, че те могат бързо да идентифицират и маркират подозрителни дейности, когато се появят, което позволява незабавна намеса, както и откриване на сложни модели и аномалии в данните [5].

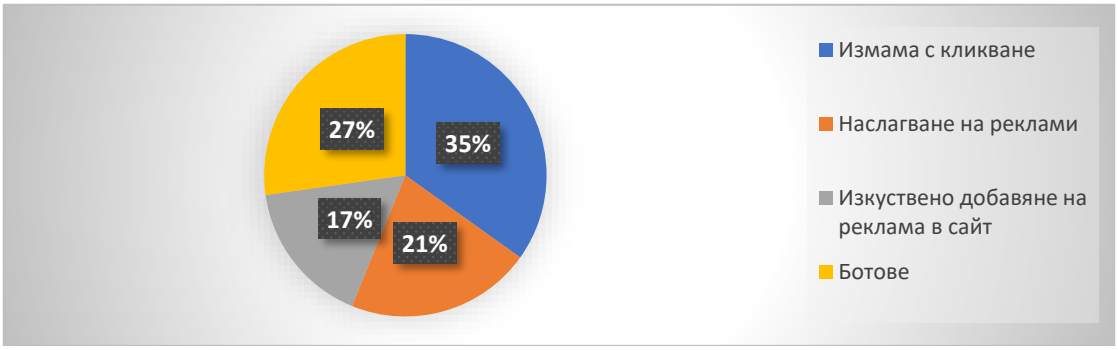
## Същност и видове рекламна измама

Разпространението на дигиталния маркетинг поставя началото на нова ера на рекламата, позволявайки на бизнеса да достигне до глобалната аудитория. Дигиталната реклама е маркетингова стратегия, която използва дигитални канали като социални медии, мобилни приложения и платформи за популяризиране на продукти, услуги или съдържание. Много от онлайн рекламите позволяват на фирмите да се насочват към конкретни демографски данни, интереси и ключови думи, като гарантират, че рекламите им достигат до аудитория, интересуваща се от техните продукти или услуги. Достигайки до правилните хора, вероятността кликванията да се превърнат в продажби или потенциални клиенти, се увеличава. Освен с всички предимства при използването на дигиталните реклами, съществуват и измами в този аспект [5].

Сферата на рекламните измами обхваща спектър от хитрости, като някои от най-разпространените методи включват внедряването на „clickbots“ – автоматизирани програми, щателно проектирани да взаимодействат с реклами, като по този начин създават разходи за рекламодателите, без да кулминира в желаното преобразуване на продажбите [7]. Друга често срещана рекламна измама се нарича „наслагване на реклами“, практика, при която отделно

рекламно пространство се продава многократно на отделен уебсайт, което води до наслагване на реклами [8]. Тази измама може да създаде илюзията за увеличен брой импресии (основен показател, който измерва колко често реклама се показва пред потенциални зрители) на рекламодателите, но въпреки това ограничава визуалното излагане на потребителите на уебсайтове само до най-горната реклама в стека. Тези импресии включват автоматизирани ботове и злонамерени участници, които генерират фалшиви изгледи и кликвания. Тези измамни дейности могат да подкопаят ефективността и целостта на дигиталните маркетингови кампании [9].

Според проучване има различни видове рекламни измами, категоризирани спрямо техния обхват, разпространение, причинени щети и т.н..В диаграмата по-долу са представени някои от най-популярните видове рекламни измами, спрямо тяхното разпространение и значение в сферата на измамите с дигитална реклама по данни от 2022г. [10].




**Диаграма 1.** Видове рекламни измами.

***Видове рекламни измами***

- **Измама с кликване** - включва повтарящи се и често автоматизирани кликвания върху онлайн реклами с намерението за изкуствено увеличаване на честотата на кликване. Както се вижда на Диаграма 1, тя е сред най-популярните измами. Тази практика не включва истински потребителски интерес, а се изпълнява, за да заблуди рекламодателите да повярват, че рекламните им кампании са по-успешни, отколкото са в действителност. Този вид измама е свързана с рекламата с плащане на клик (pay-per-click - PPC), която включва поставяне на реклами в страниците с резултати от търсачките (напр. Google Ads), където рекламодателите наддават за конкретни ключови думи и техните реклами се показват, когато потребителите търсят по тези думи. PPC е модел на онлайн реклама, при който рекламодателят плаща на платформата, където иска да се визуализира рекламата му, въз основа на броя кликове, които получава. Когато потребителите взаимодействат с реклама, като кликнат върху нея, рекламодателите се таксуват с такса. Чрез изкуствено увеличаване на разходите, намаляване на ефективността на кампанията и причиняване на неудовлетвореност за целевия бизнес, хората, реализиращи измамата, се стремят да спечелят конкурентно предимство. Това може да бъде особено пагубно по време на пиковите часове като празничните сезони, когато конкуренцията за рекламно пространство е жестока [4], [11].

- **Измама с наслагване на реклами** - практика в дигиталната рекламна индустрия, при която множество дисплейни реклами са умишлено наслоени или подредени една върху друга в рамките на едно и също рекламно разположение на уеб страница или мобилно приложение. Това може да доведе до показване на няколко реклами едновременно в едно рекламно пространство, често без знанието или съгласието на потребителя. Това се прави, за да се увеличи изкуствено броят на рекламните импресии или кликвания, като по този начин се генерират приходи за издателя (предоставящият платформата за визуализация на рекламата) и разходи за рекламодателя, дори когато потребителят може да не вижда или да не се ангажира с рекламите. Те често използват автоматизирани скриптове или ботове, за да

създават фалшиви рекламни взаимодействия, като генерират приходи, без да предоставят истинска стойност на рекламодателите [8].

- **Изкуствено добавяне на реклама в сайт** – при този вид измама рекламите се добавят в уебсайтове или уеб съдържание без съгласието на собственика на сайта или легитимната рекламна мрежа. Тези неоторизирани реклами често са с ниско качество и могат да отнемат приходи от законните издатели. Това изкуствено на реклами може да наруши потребителското изживяване чрез инжектиране на натрапчиви или неподходящи реклами, което води до отрицателно въздействие върху качеството на съдържанието [12], [13].

Осигуряването на положително потребителско изживяване е от съществено значение за запазване на доверието на потребителите. Натрапчивите реклами, бавно зареждащите се страници и неуместното насочване могат да отблъснат потребителите и да ги накарат да използват рекламни блокери, което може да намали приходите на уебсайтовете и приложенията.

### Използване на технологии с ИИ за предотвратяване на рекламни измами

В днешно време дигиталната рекламна индустрия обработва милиарди точки от данни ежедневно, което прави почти невъзможно ръчното откриване на измамни дейности поради големия обем. Изкуственият интелект (ИИ) притежава капацитета бързо да анализира огромни набори от данни, използвайки усъвършенствани алгоритми за машинно обучение, за да разпознае измамни модели и тенденции, които могат да убегнат от традиционните инструменти [5]. ИИ въплъщава идеята за разработване на машини или софтуер, които могат да имитират човешкия интелект и действия. Той включва системи, базирани на правила, които се придържат към предварително определени инструкции, както и системи, управлявани от данни, които подобряват своята производителност чрез учене от модели [14], [15].

Възможностите на ИИ са наблюдение в реално време, поведенчески анализ, откриване на ботове, разпознаване на шаблони, анализ на съдържанието, прогнозиране на измами, автоматизирано докладване и блокиране на измамен трафик. Освен това алгоритмите на ИИ могат да бъдат обучени да откриват автоматизирани ботове и да ги различават от истинския човешки трафик, използвайки сложни модели за машинно обучение, за да идентифицират поведението на ботове, характеризиращо се с бързи, повтарящи се кликвания, което се наблюдава в реално време и това помага за незабавно реагиране на аномални модели. Съществуващи решения са софтуери с използване на ИИ за откриване и блокиране на измамен трафик от ботове като White Ops, Integral Ad Science (IAS), MOAT на Oracle, Google Ads Click Fraud Protection, ClickCease и др., които откриват необичайно високи честоти на кликване, идващи от конкретен IP адрес, което показва потенциална измама с кликвания. В резултат на това се намаляват и рекламните разходи за изкуствено създаден трафик и се подобрява ефективността на рекламните кампании [16].

ИИ имат възможност за предотвратяване на измами с импресии, като ИИ следи видимостта на рекламите в реално време и преценява дали импресиите са наистина видими от реални потребители. Той постига това чрез откриване на аномалии и нередности в разположението на рекламите. Съществуващи решения са DoubleVerify, Pixalate и др. ИИ може да идентифицира кога дадена реклама се показва в скрито или малко рекламно поле, което я прави практически невидима за потребителите. В такива случаи ИИ разпознава и докладва тези случаи като измама с импресии. Този проактивен подход помага на рекламодателите да избегнат плащането за импресии, на които липсва истинска видимост, като в крайна сметка подобрява техните рекламни кампании и предотвратява финансови загуби [17].

ИИ алгоритмите са умели в откриването на рекламната измама „наслагване на реклами“. Усъвършенстваните модели за машинно обучение на ИИ разглеждат отчетените рекламни импресии спрямо действителната видимост, разпознавайки несъответствията, които показват подреждането на рекламите. Така ако една уеб страница отчете 1000 рекламни импресии, но ИИ разпознае, че само една реклама е била наистина видима за потребителите, това разкрива измамата с наслагване на реклами. Това предоставя възможност на рекламодателите да

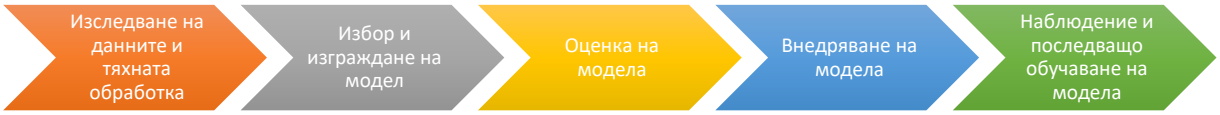
поддържат точността на кампаниите и да предотвратяват измамни практики. Съществуващо решение е системата Integral Ad Science (IAS) [18].

Рекламодателите все повече внедряват управлявани от ИИ модели, които проследяват всички действия на потребител. Така се гарантира, че правилните реклами получават заслуги за реализациите, елиминирайки възможността на измамниците да си приписват заслуги за неспечелените реализации.

**Подход за анализ на рекламни данни, свързан с откриване на измамни дейности с ИИ**

Анализът на рекламните данни е незаменим инструмент за получаване на знания, особено когато става въпрос за откриване и борба с измамни дейности. Използвайки различни технологии за извличане и анализ на данни, рекламодателите и фирмите могат да открият ценни знания, скрити в огромни масиви от данни. Справянето с нарастващото предизвикателство на рекламните измами обаче изисква нещо повече от конвенционален анализ на данни.

Поради това се предлага подход, който съчетава предварителна обработка на данни, разработване на модел, непрекъснат мониторинг и експертни познания за откриване на измамни дейности в рекламни данни с помощта с използване на Python и алгоритми на ИИ. Този подход може да бъде приспособен към всеки бизнес и множества от данни. Предложеният подход се състои от стъпките, представени на Фигура 1.




**Фигура 1.** Стъпки за подход за анализ на данни с ИИ.

На първа стъпка от Фигура 1 се изследват данните. Те могат да бъдат от различни източници, включително кликвания върху реклами, информация за устройството, IP адреси, низове на потребителски агенти и всяка подходяща контекстуална информация. Необходимо е да се познават данните, с които се работи – брой записи, типове данни, липсващи стойности и др. Данните, с които ще бъде извършено верифицирането на подхода, са от TalkingData AdTracking, пуснат на Kaggle през 2017 г. [19]. Множеството от данни съдържа над 100 000 записа за кликвания върху мобилни реклами. Целта му е да предвиди дали потребителят ще инсталира приложение, след като щракне върху реклама и дали ще разграничи измамните кликвания от легитимните. Данните са заредени и анализирани във формат DataFrame на Python. DataFrame е структура за данни, която се използва в програмния език Python, в библиотеката за анализ на данни и манипулация "pandas". DataFrame е двумерна таблица, която прилича на таблица с редове и колони. Тя позволява лесно съхранение и манипулация на данни в структуриран и организиран формат. След като данните бъдат подготвени с pandas, могат да бъдат използвани и други библиотеки като scikit-learn, TensorFlow, Keras, които са специализирани за внедряване на ИИ алгоритми.

```
import pandas as pd

data = pd.read_csv("train_sample.csv") #прочита файла
# показване на първите няколко реда от DataFrame, за да разбере структурата на данните
data.head()
```

```
#Резултат
ip   app    device   os    channel    click_time              attributed_time    is_attributed
0    87540   12      1    13    497     2017-11-07 09:30:38        NaN              0
1    105560  25      1    17    259     2017-11-07 13:40:27        NaN              0
2    101424  12      1    19    212     2017-11-07 18:05:24        NaN              0
```

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3 | 94584 | 13 | 1 | 13 | 477 | 2017-11-07 04:58:08 | NaN | 0 |
| 4 | 68413 | 12 | 1 | 1 | 178 | 2017-11-09 09:00:09 | NaN | 0 |

Наборът от данни е зареден и структурата на предоставената извадка е:

- ip: IP адресът на щракването, което е цифрово представяне, потенциално анонимно или хеширано от съображения за поверителност.
- app: Идентификационният номер на приложението за маркетинг, кодиран като цяло число.
- device: ID на типа устройство на мобилния телефон на потребителя, кодирано като цяло число.
- os: ИД на версията на операционната система на мобилния телефон на потребителя, кодиран като цяло число.
- channel: Идентификационният номер на канала на издателя на мобилна реклама, кодиран като цяло число.
- click_time: Времето на щракването в UTC.
- attributed_time: Времето, когато приложението е изтеглено, ако е изтеглено след щракване върху реклама. Това поле съдържа NaN (не число) за записи, при които не е имало изтегляне, което е обичайно за такива набори от данни, където „положителното“ събитие (изтегляне на приложение) е рядко в сравнение с броя кликвания.
- is_attributed: Целевата променлива, указваща дали приложението е изтеглено (1) или не (0) след щракване върху рекламата.

От визуализацията се вижда, че attributed_time има липсващи стойности, което има смисъл, тъй като не всички кликвания водят до изтегляния. Колоната is_attributed е двоична и ще бъде етикетът за прогнозно моделиране. Повечето от показаните тук записи не са приписани (стойността е 0), което предполага, че наборът от данни може да е дисбалансиран с много повече неизтегляния, отколкото изтегляния.

На следващо място следва да се провери общият размер на набора от данни:

```
dataset_size = data.shape
```

Оценяване на разпределението на целевата променлива is_attributed, за да се потвърди дали наборът от данни е небалансиран:

```
target_distribution = data['is_attributed'].value_counts(normalize=True)
```

```
#Резултат
((100 000, 8),
  0 0,99773
  1 0,00227
```

Наборът от данни съдържа 100 000 записа и 8 колони. Разпределението на целевата променлива is_attributed потвърждава, че наборът от данни е силно небалансиран:

- 0 (not attributed): 99,773% от записите
- 1 (attributed): 0,227% от записите

Това ниво на дисбаланс е типично при проблеми с откриването на измами, при които събитието, което има значение (в този случай изтегляне на приложение след щракване върху реклама) е рядко в сравнение с несъбитията. Такъв дисбаланс създава предизвикателства за прогнозното моделиране, тъй като повечето алгоритми за обучение предполагат равен брой примери за всеки клас или поне по-балансирано разпределение. За справяне с това могат да се използват няколко техники сред които са: повторно семплиране - чрез свръхсемплиране на малцинствения клас, недостатъчно семплиране на мнозинствения клас, или комбинация от двете; присвояване на по-висока тежест на малцинствения клас по време на обучението;

алгоритми за откриване на аномалии или ансамбъл методи, които са комбинирането на прогнозите на няколко модела.

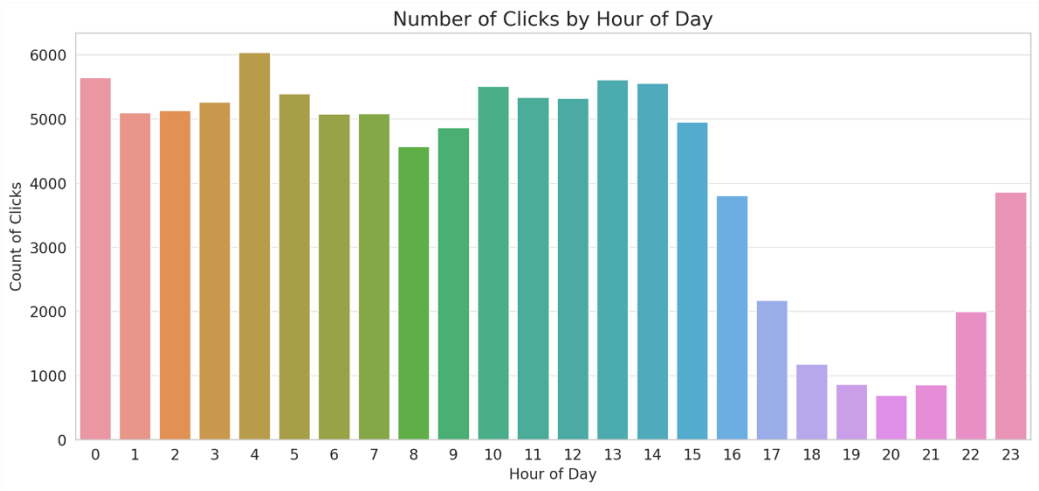
```
# проверка на типовете данни и наличието на липсващи стойности
data_info = pd.DataFrame(data.dtypes, columns=['Data Type'])
data_info['Non-Null Count'] = data.count()
data_info['Dtype'] = data.dtypes
data_info['Memory Usage'] = data.memory_usage(deep=True)
# обобщение на числовите стойности
numerical_summary = data.describe()
```

Възможно е след получаването на данните да са необходими допълнителни стъпки. Те се фокусират върху преобразуването на типа данни, анализирането на колоната „attributed_time“ и общата предварителна обработка на данни.

За извършване на анализ на данните, последните могат да бъдат визуализирани чрез графика. Кодът за нея в Python, е:

```
import matplotlib.pyplot as plt
import seaborn as sns
# задаване на стил на графиката
sns.set_style("whitegrid")
plt.figure(figsize=(14, 6))
sns.countplot(x='hour_of_day', data=data)
plt.title('Number of Clicks by Hour of Day')
plt.xlabel('Hour of Day')
plt.ylabel('Count of Clicks')
plt.show()
plt.figure(figsize=(14, 6))
# ограничава се хистограмата до времеви разлики от по-малко от 1 час (3600 секунди)
sns.histplot(data=data.loc[data['click_time_diff'] < 3600, 'click_time_diff'], bins=50, kde=False)
plt.title('Distribution of Click Time Differences (less than 1 hour)')
plt.xlabel('Time Difference (seconds)')
plt.ylabel('Frequency')
plt.show()
```

Резултатът се визуализира на Диаграма 2.




**Диаграма 2.** Графика, показваща броя на кликванията по часове.

Диаграма 2 показва броя на кликванията в различните часове на деня. От нея става ясно, че активността на кликванията варира, като в определени часове активността е по-висока. Това може да се дължи на модели на поведение на потребителите или може да показва периоди, в които измамната дейност е по-разпространена.

Тази хистограма показва честотата на разликите във времето на кликване в секунди. Има висока честота на кратки разлики във времето, която намалява бързо с увеличаването на разликата във времето. Този модел предполага, че има много случаи, при които щракванията от един и същ IP се появяват в бърза последователност, което може да е показателно за автоматизирано щракване или активност на бот.

Следващите стъпки при анализа на данните могат да бъдат:

- Инженеринг на функции: Създаване на нови функции, които биха могли да уловят по-добре моделите, свързани с измама, като функции, базирани на времето (напр. кликвания по време на деня) или функции за взаимодействие (напр. комбинации от IP и приложение).
- Визуализация на данни: График на данните за визуализиране на разпределения, връзки и потенциални аномалии.
- Предварителна обработка: Кодиране на категорични характеристики, ако е необходимо, функции за мащабиране, обработка на липсващи стойности и справяне с дисбаланса на класа.
- Избор на модел: Избор на подходящ модел или набор от модели за проблема с небалансираната класификация.
- Обучение и оценка на модела: Обучение на модела(ите) и оценяването им с помощта на подходящи показатели като площта под ROC кривата (AUC), кривата на прецизност-припомняне, матрица на объркване и др.
- Настройка на модела: Коригиране на хиперпараметрите на модела за подобряване на производителността.
- Валидиране: Използване на кръстосано валидиране, за да се гарантира, че моделът се обобщава добре към всички данни.

Предложеният подход може да бъде разширен и бъдещата работа по него е свързана с интегриране на машинно обучение в Python върху различни множества от данни за откриване на модели, зависимости и знания.

**Заключение**

В постоянно се развиващия се свят на дигиталния маркетинг едно от предизвикателствата, пред което се изправят маркетолозите, е всеобщата проблематика на измамите в рекламата. Съществуват различни видове рекламни измами, някои от които генерират огромни разходи на рекламодателя, други носят изкуствено увеличение на нивото на

ефективност от дадена кампания, както и такива които подправят домейна на легитимни сайтове с измамни. При този вид измама рекламодателите вярват, че техните реклами се показват на законни сайтове и заплащат по-висока такса за разполагане на техните реклами, когато в действителност те се озовават в измамни такива. Докато бизнесите отделят значителни бюджети за рекламни кампании, заплахата от измамните дейности продължава да подкопава ефективността и възвръщаемостта на тези усилия. Конвенционалните методи често не успяват в срок със справянето с тези проблеми и водят до значителни финансови щети. Следователно прилагането на ИИ предлага убедително решение за борба с рекламните измами. Алгоритмите с ИИ демонстрират капацитет за бърз анализ на огромни набори от данни в реално време и целесъобразно откриване на анормални модели и аномалии. Това не само възстановява доверието в дигиталната реклама, но и издига цялостния интегритет на индустрията.

## Литературни източници


[1] O. Tomyuk and O. Avdeeva, “Digital transformation of the global media market: in search for new media formats,” *Econ. Consult.*, vol. 37, Mar. 2022, doi: 10.46224/ecoc.2022.1.2.

[2] “What is Artificial Intelligence and How Does AI Work? TechTarget,” Enterprise AI. Accessed: Oct. 29, 2023. [Online]. Available: https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence

[3] A. de Lucas Ancillo and S. Gavrila Gavrila, “The Impact of Research and Development on Entrepreneurship, Innovation, Digitization and Digital transformation,” *J. Bus. Res.*, vol. 157, p. 113566, Mar. 2023, doi: 10.1016/j.jbusres.2022.113566.

[4] N. Gohil and A. Meniya, *A Survey on Online Advertising and Click fraud detection*. 2020.

[5] S. Mathur and S. Daniel, “It’s Fraud! Application of Machine Learning Techniques for Detection of Fraudulent Digital Advertising,” *Webology*, vol. 19, no. 1, pp. 2475–2490, Jan. 2022, doi: 10.14704/WEB/V19I1/WEB19166.

[6] N. Agius, “$84 billion of ad spend lost due to ad fraud in 2023,” Search Engine Land. Accessed: Oct. 28, 2023. [Online]. Available: https://searchengineland.com/ad-spend-lost-ad-fraud-2023-432610

[7] “What is click fraud? | How click bots work,” Cloudflare. Accessed: Oct. 29, 2023. [Online]. Available: https://www.cloudflare.com/learning/bots/what-is-click-fraud/

[8] “What is Ad Stacking? [+How to Prevent it].” Accessed: Oct. 29, 2023. [Online]. Available: https://edgemesh.com/blog/ad-stacking

[9] “Infographic: Where Ad Fraud is Rife,” Statista Daily Data. Accessed: Oct. 29, 2023. [Online]. Available: https://www.statista.com/chart/11347/where-ad-fraud-is-rife

[10] “Ad Fraud Statistics (2023) - Business of Apps.” Accessed: Oct. 29, 2023. [Online]. Available: https://www.businessofapps.com/ads/ad-fraud/research/ad-fraud-statistics/

[11] M. Chen, V. S. Jacob, S. Radhakrishnan, and Y. U. Ryu, “Can Payment-per-Click Induce Improvements in Click Fraud Identification Technologies?,” *Inf. Syst. Res.*, vol. 26, no. 4, pp. 754–772, Dec. 2015, doi: 10.1287/isre.2015.0598.

[12] S. Mittal, R. Gupta, M. Mohania, S. K. Gupta, M. Iwaihara, and T. Dillon, “Detecting Frauds in Online Advertising Systems,” in *E-Commerce and Web Technologies*, K. Bauknecht, B. Pröll, and H. Werthner, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2006, pp. 222–231. doi: 10.1007/11823865_23.

[13] S. Arshad, A. Kharraz, and W. Robertson, “Identifying Extension-Based Ad Injection via Fine-Grained Web Content Provenance,” in *Research in Attacks, Intrusions, and Defenses*, F. Monrose, M. Dacier, G. Blanc, and J. Garcia-Alfaro, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016, pp. 415–436. doi: 10.1007/978-3-319-45719-2_19.

[14] N. Makkineni, A. Ciripuram, S. N, S. Subhani, and V. Kakulapati, “Fraud Detection of AD Clicks Using Machine Learning Techniques.” Rochester, NY, Jun. 20, 2023. doi: 10.2139/ssrn.4486834.

[15] I. Buch and M. Thakkar, “AI in Advertising,” 2021. doi: 10.13140/RG.2.2.19912.24323.

[16] M. Aljabri and R. M. A. Mohammad, “Click fraud detection for online advertising using machine learning,” *Egypt. Inform. J.*, vol. 24, no. 2, pp. 341–350, Jul. 2023, doi: 10.1016/j.eij.2023.05.006.

[17] A. Siu, “How the growth of click and impression farming are getting worse with AI,” Digiday. Accessed: Oct. 30, 2023. [Online]. Available: https://digiday.com/media-buying/the-growth-of-click-and-impression-farming-are-getting-worse-with-ai/
[18] J. R. Alzghoul, E. E. Abdallah, and A. S. Al-khawaldeh, “Fraud in Online Classified Ads: Strategies, Risks, and Detection Methods: A Survey,” *J. Appl. Secur. Res.*, vol. 0, no. 0, pp. 1–25, 2022, doi: 10.1080/19361610.2022.2124328.
[19] “TalkingData AdTracking Fraud Detection Challenge.” Accessed: Nov. 04, 2023. [Online]. Available: https://kaggle.com/competitions/talkingdata-adtracking-fraud-detection.